DOCUMENT RESUME

ED 468 714 TM 034 423

AUTHOR Meehan, Merrill L.; Cowley, Kimberly S.; Wiersma, William; Orletsky, Sandra R.; Sattes, Beth D.; Walsh, Jackie A.

TITLE Development and Testing of an Instrument Measuring a School Faculty's Commitment to Continuous Learning and Improvement.

INSTITUTION AEL, Inc., Charleston, WV.

SPONS AGENCY Office of Educational Research and Improvement (ED), Washington, DC.

PUB DATE 2002-07-00

NOTE 28p.; Paper presented at the Annual Meeting of the CREATE National Evaluation Institute (11th, Boise, ID, July 25-27, 2002).

CONTRACT ED-01-CO-0016

PUB TYPE Reports - Research (143) -- Speeches/Meeting Papers (150)

EDRS PRICE EDRS Price MF01/PC02 Plus Postage.

DESCRIPTORS *Educational Improvement; Elementary Secondary Education; *Measurement Techniques; Reliability; Surveys; *Teacher Attitudes; Teachers; Test Construction; Validity

IDENTIFIERS AEL Regional Exchange; *Commitment; QUEST Program

ABSTRACT

As part of its school improvement effort, AEL, a regional education laboratory, developed the Continuous School Improvement Questionnaire (AEL CSIQ). Staff from the AEL Quest schools program drafted a 65-item questionnaire to help measure and assess the efforts of the project team in their work with the 18 schools in the Quest network. These items embodied the six major concepts of the Quest school improvement program. After review and refinement, 42-item and 117-item versions of the scale were developed and combined for pilot testing. Pilot and field tests provided much information about the AEL CSIQ and its measurement. With data from more than 3,800 educators in 132 schools, the statistics generated have high stability and precision. The testing of the AEL CSIQ in the field fostered its development into an excellent and usable 6-item inventory. The CSIQ was reduced from 147 items with unequal numbers of items in subscales to 10 items in each subscale. The AEL CSIQ was found to be very sound technically. Internal consistency reliability was sound for all subscales, as was stability reliability. The instrument showed satisfactory concurrent validity with another school climate instrument. The items had face validity based on content, and the factor analysis provided strong support for the construct validity of the entire inventory as reflected by the subscales. (Contains 12 tables, 1 figure, and 5 references.) (SLD)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 468 714

Development and Testing of an Instrument Measuring a School Faculty's Commitment to Continuous Learning and Improvement

Merrill L. Meehan, AEL
Kimberly S. Cowley, AEL
William Wiersma, University of Toledo
Sandra R. Orletsky, AEL
Beth D. Sattes, AEL
Jackie A. Walsh, AEL

AEL
P.O. Box 1348
Charleston, WV 25325

Paper presented at the 11th Annual CREATE National Evaluation Institute
Boise, Idaho
July 25-27, 2002

TM034423

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

BEST COPY AVAILABLE

AEL is a catalyst for schools and communities to build lifelong learning systems that harness resources, research, and practical wisdom. AEL serves as the Regional Educational Laboratory for Kentucky, Tennessee, Virginia, and West Virginia. For these same four states, it operates the Eisenhower Regional Consortium for Mathematics and Science Education. The Region IV Comprehensive Center at AEL serves North and South Carolina in addition to these states. AEL also operates the ERIC Clearinghouse on Rural Education and Small Schools.

Information about AEL projects, programs, and services is available by writing or calling AEL.

Post Office Box 1348
Charleston, West Virginia 25325-1348
304-347-0400
800-624-9120
304-347-0487 (fax)
aelinfo@ael.org
http://www.ael.org

© 2002 by AEL

This publication is based on work sponsored wholly or in part by the Office of Educational Research and Improvement, U.S. Department of Education, under contract number ED-01-CO-0016. Its contents do not necessarily reflect the views of OERI, the Department, or any other agency of the U.S. government.

AEL is an Equal Opportunity/Affirmative Action Employer.

TABLE OF CONTENTS

INTRODUCTION 1

METHODS 3
- Instrument 3
- Sample 4
- Data Analyses 5

FINDINGS 7
- Pilot Test 7
  - Internal Consistency Reliabilities 7
  - Item-to-Subscale Score Correlations 8
  - Relative Positioning of the Scales 9
  - Correlations Among the Subscales and Total Score 12
  - Concurrent Validity 13
- Field Tests 14
  - Stability Reliability 14
  - Construct Validity 15
  - Gender Effects 19

CONCLUSIONS 21
- Pilot Test 21
- Field Tests 21

REFERENCES 23

LIST OF TABLES

1: Cronbach Alpha Reliability Estimates for AEL CSIQ Subscales and Total Scale by Number of Items 7

2: Range of Item-to-Subscale Correlations for AEL CSIQ Subscales and Total Scale by Number of Items 8

3: Means for AEL CSIQ Subscales and Total Score by Number of Items 9

4: Location of the Mean in Percent of Distance from Minimum to Maximum Scores for AEL CSIQ Subscales and Total Score, by Number of Items 10

5: Standard Deviations for AEL CSIQ Subscale and Total Score by Number of Items 11

6: Coefficients of Variation (Expressed as Percents) for AEL CSIQ Subscales and Total Score, by Number of Items 12

7: Correlation Matrix for AEL CSIQ 12-Item Subscales and Total Score 13

8: Frequencies (N) and Correlation Coefficients (r) for the Test-Retest of the AEL CSIQ Subscales and Total Score 14

9: AEL CSIQ Factor Loadings by Item Number and Factor Subscale, Elementary School Group 16

10: AEL CSIQ Factor Loadings by Item Number and Factor Subscale, High School Group 17

11: Factor Loadings by Item Number and Factor Subscale, Combined Group 18

12: Frequencies (N) and Means (M) for AEL CSIQ Subscales and Total Score by Gender 20

LIST OF FIGURES

1: AEL's Framework for Transforming Low-Performing Schools into High-Performing Learning Communities 2

INTRODUCTION

Attempts to improve schools seem to be as old as schools themselves. School reform, school improvement, and school restructuring are some of the more recent efforts to improve our nation's schools. These school improvement efforts have met with various degrees of success. Many of these programs have been funded by federal and state funds.

The staff of a school planning an improvement program faces two major challenges: (1) "getting a handle on" how to begin and (2) keeping the improvement process going. To address the first challenge, a logical starting point is measuring and assessing a school on a "scale" of improvement. To do this, schools need a conceptual framework that outlines and portrays the dimensions of school improvement. The school's instructional program and the elements that impact it--such as home and family--include many interrelated items, some wholly or partially under the direct control of the school and others over which the school exercises no control.

The ultimate goal of any school improvement effort is to improve student achievement. Although there may be other goals in school improvement, they are secondary to the improvement of student achievement goal. This goal has received increased prominence due to recent emphases to proficiency and performance testing.

AEL, in its role as a regional educational laboratory, has been committed to research on school improvement efforts since 1966. One of AEL's more recent, successful projects was Quest (1996-2000), a network of schools located in AEL's regional states of Kentucky, Tennessee, Virginia, and West Virginia. Quest schools were dedicated to building learning communities that support high levels of student and adult performance. The Quest Network for Quality Learning communities emphasized six components, or dimensions, essential for student and adult learning. These dimensions evolved into the conceptual framework that undergirds the *AEL Continuous School Improvement Questionnaire* (AEL CSIQ). During their collaboration with AEL, Quest schools contributed to the AEL CSIQ research base and participated in the pilot test (Howley-Rowe, 2000a, 2000b, 2000c, 2000d).

AEL's basis for school improvement is embodied in its research-based framework. AEL's Framework for Transforming Low-Performing Schools into High-Performing Learning Communities underlies the AEL CSIQ. See Figure 1 for a graphic depiction of AEL's framework for the AEL CSIQ and other AEL products and services designed to help in the development of high-performing learning communities. The very center of the framework is the circle representing improved student performance, indicating that all school improvement efforts ultimately are directed to this central goal. The six dimensions of high-performing learning communities are depicted as circling the increased student performance goal. They are: shared leadership, effective teaching, school/family/community connections, purposeful student assessment, shared goals for learning, and learning culture. The arrows outside the six dimensions indicate the dynamic nature of the framework. The dimensions are interactive and they require continuous action on the part of the school community.

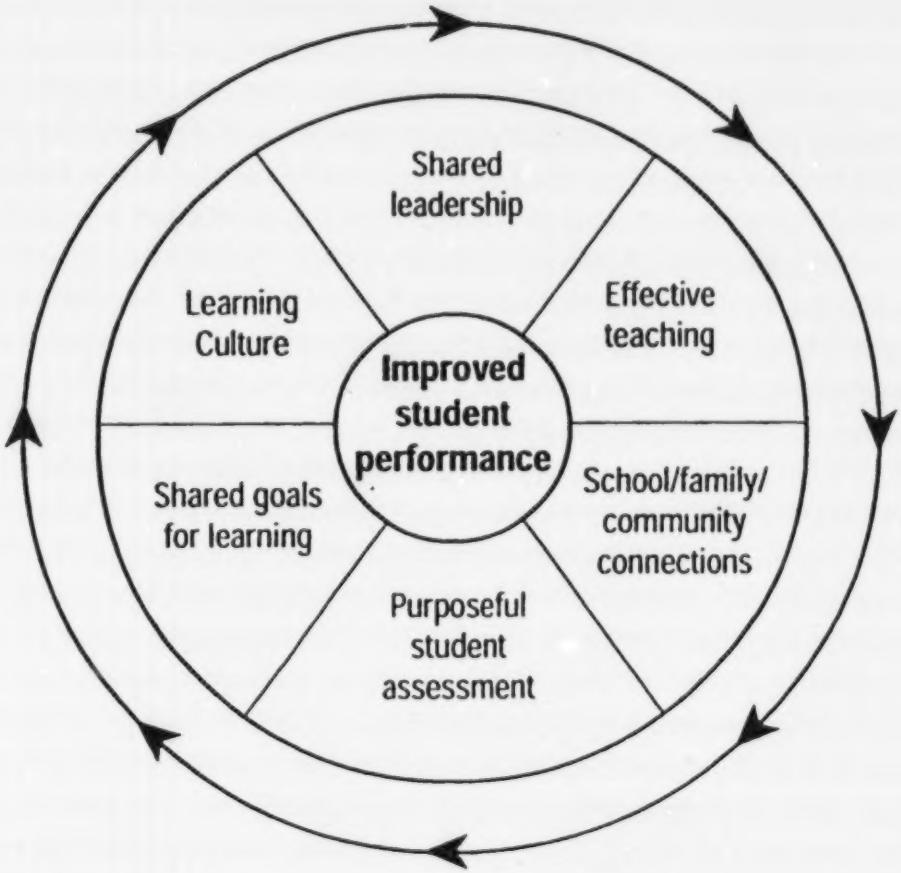


Figure 1. AEL's Framework for Transforming Low-Performing Schools into High-Performing Learning Communities

METHODS

Instrument

The development and testing of the AEL CSIQ is the subject of this paper. Therefore, this section will explain the drafting of the items that were tested and analyzed several times.

The genesis for the AEL CSIQ was in AEL's Quest research project, which was a four-year effort to form and operate a network of schools interested in working together with the goal of learning how to improve together as a faculty, school, and community. Quest project staff built their project around six overarching concepts that guided their multi-year work and activities. An instrument was needed to help measure and assess the efforts of the project team in their work with the 18 schools in the Quest network. Project staff drafted an initial set of 65 questionnaire item statements for a survey of school staff. In this initial set of items, there were between 10 to 12 per each of six major concepts in the Quest model. At this stage, the response option for each item was a 10-point scale of low to high. Project staff submitted these first 65-items to AEL research and evaluation staff for review and processing.

AEL research and evaluation staff reviewed the initial set of 65 items and concluded that, although they reflected the content of the Quest project well, they possessed technical deficiencies. The most common technical deficiency in the set of 65 items was that many addressed multiple concepts in the same item stem. Many addressed two, three, and even four subjects in one item statement. The research and evaluation staff split up these multiple-subject statements, which resulted in an expanded set of 117 item statements. Also, the suggestion was made at this stage to change the response option from a 10-point scale to a 5- to 7-point scale.

While the research and evaluation staff were converting the initial set of 65 items into 117 items, Quest project staff developed another version of their instrument that adopted the suggested 6-point response option scale, but retained many of the multiple-concept item statements from the initial survey and also dropped some items. The version of the instrument Quest staff wanted to pilot test now included 42 items. The version advocated by the research and evaluation staff included 117, of which 12 were also in the 42-item version. AEL management decided to test all the different items in the pilot test and make reductions in items based on the resulting data from school practitioners who completed the draft of the instrument.

To summarize, 30 items from the 42-item project staff instrument were added to the 117 items that the research and evaluation staff developed from the initial set of 65 items. Thus, the pilot test version of the AEL CSIQ contained 147 items. The number of items for each of the six subscales ranged from 20 to 27. The response option for all 147 items was a 6-point scale from "Is not present" to "Is present to a high degree."

Subsequent versions of the AEL CSIQ instrument were based on the data from the pilot and field tests and will be explained in the Findings sections.

Sample

The AEL CSIQ was designed to be a general inventory that can be used with educators at any school level from kindergarten through 12th grade. The instrument, or, more correctly, variations of the original instrument were tested by AEL staff three times: one pilot test and two field tests.

The pilot test of the AEL CSIQ was conducted in the spring of 2000. Here, the 147-item version was administered to educators in 28 different schools in AEL's four-state region of Kentucky, Tennessee, Virginia, and West Virginia. Since the major purpose of the pilot test was to collect data for use in reducing the length of the draft version of the AEL CSIQ, there was no concerted effort to collect and report on respondents' schools. The pilot test data were not cut by any school-level variables. Respondents for the pilot test were recruited from contacts and participants in various AEL projects and activities at that time. A total of 274 usable AEL CSIQ instruments were returned to AEL for analyses in the pilot test. Most of those respondents were teachers, although other school educators, such as principals, assistant principals, counselors, media specialists, etc. were encouraged to complete and return the instruments. The pilot-test version consisted of the 147 items, in six subscales, printed on the front side of four sheets of white paper and stapled in the upper left corner.

Based on the pilot-test results, the original 147-item AEL CSIQ was reduced to 12 items for each of the six subscales or 72 items. The first field test of the 72-item version of the instrument was administered in the fall of 2000 to a total of 2,093 educators, mostly teachers. Since the purposes of the field tests were to determine performance levels for a larger base of educators and also to investigate possible differences by the school levels, respondents' school information was sought. In the first field test, respondents were from 79 different schools. Again, AEL staff recruited volunteers within the four-state region to complete the shortened instrument and, again, in AEL's four-state region. This same 72-item version of the AEL CSIQ was also administered to volunteer educators in 25 schools in Tennessee in the winter of 2000. Even though the timing of this administration of the instrument was several months after the first administration of the 72-item version, this was considered part of the first field test by virtue of the number of items in the instrument.

Then, to make the AEL CSIQ more convenient to use, but still retain satisfactory internal consistency reliabilities for the six subscales and the total score, the decision was made to reduce its length to 60 items, 10 per subscale. Also, the decision was made to place all the items in random order, rather than grouped by subscale name, as in the pilot test and the first field test. This 60-item version of the AEL CSIQ was administered to the full faculties of 75 schools in Tennessee that were participating in a project with AEL staff. In effect, then, this second field test consisted of a convenience sample of educators, whereas the respondents in the pilot test and first field test either volunteered themselves or were volunteered by others to complete the AEL CSIQ.

BEST COPY AVAILABLE

Data Analyses

Various types of data analyses were conducted for the pilot and field tests of the AEL CSIQ. These analyses will be described in this subsection.

For the pilot test, the first analysis was of the quality of the data entry for the returned AEL CSIQs. This step was necessary because the instrument was not printed in a manner compatible with optical scanning of data entry--the data were entered into the SPSS software program by hand. The quality control check for out-of-range entries located 22 instances out of a possible 43,018 entries. This was an error rate of 0.05%. These errors were corrected. Thirty-five questionnaires were selected at random and checked by matching responses to data entry. Forty-four data entry errors were found for 5,495 entries, an error rate of 0.08%. These also were corrected.

Knowing that the pilot test version of the AEL CSIQ consisted of both too many and deficient items, the first step in the analyses was to locate the technically deficient items, flag them, and reduce the subscales down to a more manageable number. In this step, 41 deficient items were identified. Then, the effects of eliminating these items were determined by computing Cronbach alpha internal consistency reliabilities for the original length subscales, the subscales reduced to 15 items each, and the subscales reduced to 12 items each. Also, the item-to-subscale score correlations were computed. Items with the lowest item-to-subscale correlations were dropped each time the instrument was shortened.

The relative positioning on the measurement scales was the next target for analysis. Here, the aim was to assess the spread of the scores for each subscale. This was accomplished by computing and displaying the means for the six subscales and the total score for the three different number of items per subscale (original number, 15, and 12 items). Then, the locations of the means in terms of the percent of distance from the minimum and maximum scores were computed. Also, the standard deviations of the AEL CSIQ subscale and total scores were computed. The coefficients of variation, expressed as a percent, for the six subscales and the total score were computed. Last, in this phase of analysis, the correlations among the subscales and the total score were computed.

As a measure of concurrent validity, the *School Climate Questionnaire* (Manning, Curtis, & McMillen, 1996) was administered to the 274 educators in the pilot test of the AEL CSIQ. The *School Climate Questionnaire* is an instrument measuring the perceptions of the respondents regarding their school. It contains 10 rating scale items, each with a 20-point response option. AEL staff determined that it was closest to what the AEL CSIQ measures, after reviewing several possible instruments. The correlation between the AEL CSIQ total score (all three versions, i.e., 147 items, 90 items, and 72 items) to the *School Climate Questionnaire* total score were computed as measures of concurrent validity.

For the field tests, the data analyses began by combining all of the AEL CSIQ files together. This was necessary because some of the field-test administrations were with the 72-item version of the instrument, while the last administration was with the 60-item version. So, the first step in combining the files was to eliminate the same two items from the 72-item version

file as were dropped to create the 60-item version. The second step before combining the field-test files was to drop those schools with less than 10 AEL CSIQ respondents. Recall, school staffs volunteered to complete the instrument, but there was no incentive nor reward for doing so. Thus, completion of the instrument varied by school. In some schools, only a few educators on staff returned completed copies, while in other schools, there was near total completion rates. After the file transformations were completed, the resulting database for the field-test analyses included 3,821 cases in 132 schools.

To obtain a measure of stability reliability—to supplement the internal consistency reliabilities computed—approximately 300 educators in one of the field tests completed the AEL CSIQ twice with about a three-week interval between administrations. The educators in the test-retest effort were located in 20 schools across AEL's four states including seven elementary, three middle, six high, and four middle/high configurations. Correlation coefficients between the two administrations for the AEL CSIQ subscale and total scores were computed.

As another measure of validity, a factor analysis of the AEL CSIQ field-test database was conducted. The factor analysis computed was a principal axis analysis with a Varimax rotation. This factor analysis was computed on the item scores in the database, not the scores of the subscales. A .30 blankout was used to help in the inspection and interpretation of the results. The scree plots and the factor loadings were also generated and used in the interpretation. Three different factor analyses were conducted on the field tests database: the elementary group, the high school group, and the elementary and high school groups combined. The elementary group included scores for about 1,750 educators in 81 schools and the high school group included scores for about 680 educators in 17 schools. There were insufficient numbers of middle school educators in the database to merit a factor analysis.

Finally, as part of the field test of the AEL CSIQ, each respondent was asked a few demographic questions, including gender. The AEL research team recognized that the majority of the schools in the database were elementary schools and that the majority of the respondents in those schools were females; nonetheless, gender differences, if any, were important to locate. Thus, independent t-tests were computed on the six subscale scores and the total score, testing the difference between male and female respondents.

FINDINGS

Pilot Test

Internal Consistency Reliabilities

Table 1 presents the Cronbach alpha internal consistency reliability estimates for the six subscales and the total score. Per the pilot-test analysis plan, deficient items were located within each subscale in order to reduce them to a more convenient and technically-correct version. Thus, the data in Table 1 are provided for (1) the original number of items per subscale, (2) the 15-item version, and (3) the 12-item version.

Table 1
Cronbach Alpha Reliability Estimates for AEL CSIQ Subscales and Total Scale by Number of Items

Subscale	Original - S 20 to 27 Items	15-Item Subscales	12-Item Subscales
S_1	.94	.90	.89
S_2	.97	.95	.95
S_3	.98	.97	.97
S_4	.96	.96	.95
S_5	.95	.93	.92
S_6	.97	.96	.95
Total	.99	.99	.98

S_1: Learning Culture
S_2: School/Family/Community Connections
S_3: Shared Leadership
S_4: Shared Goals for Learning
S_5: Purposeful Student Assessment
S_6: Effective Teaching

Inspection of the alpha reliability coefficients for the pilot test shows them to be consistently high for the six subscales and the total score. For inventories of this nature, reliability coefficients around .80 are satisfactory. Only one subscale S_1 had a reliability coefficient of less than .90 and that one was .89 for the 12-item S_1 subscale.

Alpha reliability estimates for the original-length subscales, those with 20 to 27 items, ranged from .94 to .98, with the total scale having a reliability of .99. When all the subscales

were reduced to 15 items by dropping the technically deficient items and those with the lowest item-to-subscale correlation, the subscale reliabilities ranged from .90 to .97, with the reliability of the total scale remaining at .99. All the items on the 15-item subscales were well constructed technically, so three more items were eliminated (again, those with the lowest item-to-subscale correlations). The reliability estimates for the 12-item subscales ranged from .89 to .97 and the total 72-item scale was .98. Thus, the data in Table 1 show that there were no noteworthy reductions in reliability estimates for either the 15-item reduction or the 12-item reduction in the pilot test.

Item-to-Subscale Correlations

Table 2 displays the range of item-to-subscale correlations for the AEL CSIQ subscales and the total scale score. Again, the three columns of data are for the three different numbers of items in the subscales (original, 15 items, and 12 items). The ranges of the correlations essentially were unaffected by the reductions in the numbers of subscale items. For some of the subscales, the upper range of the correlations increased very slightly. For example, S_3 had the highest range for the 12-item subscale, over the longer versions of that subscale.

Table 2
Range of Item-to-Subscale Correlations for AEL CSIQ Subscales and Total Scale by Number of Items

Scale	Original – S 20 to 27 Items	15-Item Subscales	12-Item Subscales
S_1	.51 to .71	.51 to .65	.51 to 65
S_2	.68 to .83	.64 to .81	.69 to .81
S_3	.66 to .88	.65 to .89	.79 to .90
S_4	.62 to .84	.61 to .84	.62 to .84
S_5	.57 to .72	.52 to .71	.57 to .72
S_6	.71 to .81	.69 to .82	.73 to .81
Total	.46 to .82.	.44 to .80	.45 to .80

S_1: Learning Culture
S_2: School/Family/Community Connections
S_3: Shared Leadership
S_4: Shared Goals for Learning
S_5: Purposeful Student Assessment
S_6: Effective Teaching

Relative Positioning of the Scales

If in fact, there are variations in whatever is measured, it is desirable to obtain a spread in the scores. Also, scores should at least get on the scale of measurement, but they should not be so high that practically everybody is at the ceiling of the test or inventory.

Table 3 contains the means for the AEL CSIQ subscales and the total score for the three different numbers of items in the pilot test. The original version had varying numbers of items per subscale so the maximum possible scores are given in parentheses in the table. The minimum score is 15 for the 15-item subscale and 12 for the 12-item subscale. The minimum scores for the original version vary, but for any subscale, it equals the number of items.

Table 3

Means for AEL CSIQ Subscales and Total Score by Number of Items

Subscale	Original[1]	15-Item Max 90 for Subscales	12-Item Max 72 for Subscales
S_1	88.6 (138)	58.7	47.0
S_2	98.7 (156)	56.3	44.6
S_3	97.8 (156)	56.3	45.4
S_4	78.8 (120)	59.1	47.2
S_5	103.5 (150)	61.9	49.6
S_6	119.0 (162)	67.7	54.1
Total	581.8 (882)	360.9 (540)	288.6 (432)

[1]Maximum scores given in ( ) for varying number of items.

S_1: Learning Culture
S_2: School/Family/Community Connections
S_3: Shared Leadership
S_4: Shared Goals for Learning
S_5: Purposeful Student Assessment
S_6: Effective Teaching

Data in Table 3 show that S_6 consistently has the higher subscale mean score in both 15-item and 12-item versions. In fact, for the 12-item version, S_6 was the only subscale mean score above 50 points (maximum equal 72 points). For the 15-item version, two subscales (S_6 and S_5) had mean scores in the 60s (maximum equal 90 points). On the low end, in both reduced versions, S_2 had the lowest mean score, although for the 15-item version, S_3 tied for the lowest mean score.

In order to check on the location of the means and the consistency of the location across the different numbered items of the AEL CSIQ, it is useful to consider the percentage of the distance from the minimum to maximum score that locates the mean. To illustrate how these percentages are determined, we will use the 15-item version of S_1:

15-item minimum score = 15, maximum score = 90,
a 75-point distance; the mean score of 58.7 was
43.7 points in from the minimum

$$\frac{43.7}{75} = 58\%$$

Table 4 shows the location of the mean in percent of distance from the minimum to maximum scores for the AEL CSIQ subscales and the total scores for all three versions. The data in Table 4 shows the consistency of location of the mean across the varying number of items. For two subscales (S_4 and S_5), there was no change at all; further, the greatest change was only two percentage points and that occurred in two subscales (S_2 and S_6). Thus, reducing the number of items in the AEL CSIQ subscale essentially had no effect on the location of the means on the scales of measurement.

Table 4
Location of the Mean in Percent of Distance from Minimum to Maximum Scores for AEL CSIQ Subscales and Total Score by Number of Items

Scale	Original	15-Item Subscales	12-Item Subscales
S_1	57	58	58
S_2	56	55	54
S_3	55	55	56
S_4	59	59	59
S_5	63	63	63
S_6	68	70	70
Total	59	60	60

S_1: Learning Culture
S_2: School/Family/Community Connections
S_3: Shared Leadership
S_4: Shared Goals for Learning
S_5: Purposeful Student Assessment
S_6: Effective Teaching

Another point can be made about the data in Table 4. The AEL CSIQ respondents clearly were on the scale of measurement, but they were 54% to 70% above the minimum. So, there is room for change in either direction.

Table 5 displays the standard deviations for AEL CSIQ subscale and total scores for all three versions. These data show the dispersion or variability in the scores. As expected, the original version with the longer subscales tends to have the larger standard deviations or more variability. The standard deviations for the 15-item and 12-item subscales are large enough to show a desirable amount of variation.

Table 5
Standard Deviations for AEL CSIQ Subscales and Total Score by Number of Items

Subscale	Original	15-Item Subscales	12-Item Subscales
S_1	17.7	11.3	9.3
S_2	24.5	14.9	12.2
S_3	28.2	17.2	14.6
S_4	19.5	14.9	12.1
S_5	18.7	12.1	10.1
S_6	20.5	11.8	9.6
Total	117.7	72.0	58.5

S_1: Learning Culture
S_2: School/Family/Community Connections
S_3: Shared Leadership
S_4: Shared Goals for Learning
S_5: Purposeful Student Assessment
S_6: Effective Teaching

The coefficient of variation is the standard deviation divided by the mean. This can be thought of as a "relative standard deviation." It is a unitless number, comparable across various means and scales. Too, the coefficient of variation can be expressed as a percent. In this case, it is the standard deviation expressed as a percent of the mean. That is, the ratio S/x is converted to a percent.

Table 6 presents the coefficients of variation, expressed as percents, for the AEL CSIQ subscales and total score for all three versions. The displayed coefficients of variation are quite consistent, especially for individual subscales and the total score across numbers of items (versions). The differences in the coefficients of variation ranged from a low of 0.2% across the three versions (total) to a high of 3.4% (S_3).

Table 6
Coefficients of Variation (Expressed as Percents) for AEL CSIQ Subscales and Total Score by Number of Items

Scale	Original	15-Item Subscales	12-Item Subscales
S_1	20.0	19.3	19.8
S_2	24.8	26.5	27.4
S_3	28.8	30.6	32.2
S_4	24.7	25.2	25.6
S_5	18.1	19.5	20.4
S_6	17.2	17.4	17.7
Total	20.2	20.0	20.2

S_1: Learning Culture
S_2: School/Family/Community Connections
S_3: Shared Leadership
S_4: Shared Goals for Learning
S_5: Purposeful Student Assessment
S_6: Effective Teaching

Correlations Among the Subscales and Total Score

Three correlation matrices were generated for the subscale and total scores on the AEL CSIQ; one for each version of the instrument. The correlation patterns were very similar across the three versions of the instrument. Therefore, only the correlation matrix for the 12-item subscales is presented here.

Table 7 is the correlation matrix for the 12-item subscales and total score of the AEL CSIQ administered in the pilot test. The correlations among the subscales range from .45 to .81. Ten of the 15 correlations among the subscales were in the .60s and .70s and two correlations were in the .80s. Thus, there are moderate relationships among the scores of the subscales. With the moderate correlationships among subscale scores and the total score being the sum of the subscale scores, relatively high correlations between subscale and total scores would be expected. This was the case, as shown in Table 7. Those correlations ranged from .75 to .92, with only one correlation below .84.

Table 7
Correlation Matrix for AEL CSIQ 12-Item Subscales and Total Score

	S_2	S_3	S_4	S_5	S_6	Total
S_1	.72	.59	.68	.78	.64	.84
S_2		.81	.79	.70	.54	.90
S_3			.77	.63	.45	.85
S_4				.81	.61	.92
S_5					.72	.89
S_6						.75

S_1: Learning Culture
S_2: School/Family/Community Connections
S_3: Shared Leadership
S_4: Shared Goals for Learning
S_5: Purposeful Student Assessment
S_6: Effective Teaching

Concurrent Validity

Concurrent with the pilot-test administration of the AEL CSIQ, the *School Climate Questionnaire* (SCQ; Manning, Curtis, & McMillen, 1996) was administered to a volunteer group of 274 educators. The SCQ is a one-page questionnaire soliciting respondents' perceptions about their school. The SCQ contains 10 rating items, each having a rating scale of 20 points. The internal consistency reliability for the SCQ in the pilot test was .97. The ten items were judged to relate to school improvement. The correlation of scores on the SCQ with scores on the AEL CSIQ serves as a measure of the concurrrent validity for the AEL CSIQ.

The correlations between the AEL CSIQ total scores, for all three versions, and the total SCQ scores were computed. These correlations were:

Original 147 – Item Total = .76
90 – Item Total = .75
72 – Item Total = .76

Thus, the concurrent validity measure was very stable across the different lengths of the AEL CSIQ and the correlations indicate satisfactory concurrent validity.

Field Tests

Stability Reliability

Throughout the testing of the AEL CSIQ, even as the numbers of items per subscale were reduced twice, the internal consistency reliability of the subscale and total scale scores remained high. In order to obtain a measure of stability reliability, approximately 300 educators in the field test were measured twice, allowing about a three-week interval between the two administrations of the instrument.

Table 8 presents the results of the test-retest administration of the AEL CSIQ. Shown in the table are the frequencies and correlation coefficients for the two administrations of the instrument. The frequencies and correlations are for the six subscales and the total score. The test-retest correlations ranged from .66 to .81, showing two results: (1) considerable stability across time for all seven scores and (2) similar levels of stability across the AEL CSIQ scores.

Table 8
Frequencies (N) and Correlation Coefficients (r) for Test-Retest of the AEL CSIQ Subscales and Total Score

Subscale	N	r
S_1	306	.70
S_2	305	.75
S_3	299	.81
S_4	293	.76
S_5	296	.70
S_6	303	.66
Total	257	.80

S_1: Learning Culture
S_2: School/Family/Community Connections
S_3: Shared Leadership
S_4: Shared Goals for Learning
S_5: Purposeful Student Assessment
S_6: Effective Teaching

The Pearson product-moment correlation coefficient is a measure of the relative positioning of the scores being correlated and the magnitude of the difference between these

scores on the variables. An inspection of the test-retest mean scores (not shown in Table 8) showed very little differences—less than 1.0—in the subscale means. For most of the subscales, the retest mean was the greater, however slightly, of the pair.

Construct Validity

Factor analysis is one of the chief methods used to establish construct validity for an instrument. Factor analysis is a statistical technique that generates, from items' responses, artificial variables (called factors) representing one or more constructs measured by the entire inventory or test, in this case the 60 items in the field-test version of the AEL CSIQ. Although factors are artificial variables, they are defined or described in terms of the variables (60 items) on which they are based. Factor loadings, which are correlations between the scores on individual items and factors, serve this purpose. Thus, a high positive factor loading indicates that the item contributes extensively to the composition of the factor.

The desirable outcome of factor analysis is to have as many noteworthy factors as there are logical constructs underlying the concept being studied. For the AEL CSIQ, it would be desirable to have the 60 items form six factors. The commonly used criterion for retaining extracted factors is to retain those with an eigenvalue of 1.0 or greater. That criterion was employed in the three AEL CSIQ factor analyses with one exception. The sixth factor extracted from the high school group had an eigenvalue of .924, but it was retained because it was high and the factor fit subscale S_5.

Results of the three factor analyses are displayed in the following Tables 9, 10, and 11. Instead of displaying 360 factor loadings per table, each contains 60 loadings. The accepted criterion of reporting loadings equal to or greater than .30 was used in constructing these tables. The desirable outcome is for the items of a subscale to load heavily (i.e., have high correlations with) one factor and to have zero or low loadings on other factors. This was the outcome for all three factor analyses. The results displayed in Tables 9, 10 and 11 show the 10 items and the factor most closely representing the construct measured by the subscale. For example, consider factor 1 in Table 9. Here, S_6 loads heavily on this factor and the 10 factor loadings range from .60 to .78, certainly substantial loadings. This (S_6) is the Effective Teaching subscale and its related construct most heavily underlies continuous improvement in the elementary schools.

Tables 9, 10, and 11 show interesting similarities and differences from the factor analyses. Although there were occasional loadings greater than .30 for items from a subscale other than the one associated with a factor, these were few and the loadings tended to be only slightly above .30. So, these factor analyses of the AEL CSIQ were exceptionally "clear." That is, each subscale was closely associated with only one factor, indicating that the AEL CSIQ has six underlying constructs, each measured quite well by a subscale.

Table 9
AEL CSIQ Factor Loadings by Item Number and Factor Subscale, Elementary School Group

	Factor Number and Corresponding Subscale					
Item No.	$1/S_6$	$2/S_3$	$3/S_4$	$4/S_2$	$5/S_1$	$6/S_5$
1	.60	.75	.46	.40	.56	.38
2	.73	.77	.64	.65	.53	.44
3	.70	.74	.63	.75	.72	.42
4	.69	.82	.68	.74	.58	.53
5	.70	.85	.69	.61	.72	.55
6	.72	.68	.64	.66	.36	.61
7	.78	.81	.57	.41	.30	.57
8	.74	.72	.64	.39	.41	-.57
9	.77	.75	.54	.34	.45	.59
10	.77	.56	.56	.39	.33	.38

S_1: Learning Culture
S_2: School/Family/Community Connections
S_3: Shared Leadership
S_4: Shared Goals for Learning
S_5: Purposeful Student Assessment
S_6: Effective Teaching

Table 10
AEL CSIQ Factor Loadings by Item Number and Factor Subscale, High School Group

	Factor Number and Corresponding Subscale					
Item No.	$1/S_3$	$2/S_6$	$3/S_4$	$4/S_2$	$5/S_1$	$6/S_5$
1	.73	.62	.42	.50	.53	.34
2	.74	.73	.57	.64	.58	.38
3	.78	.72	.60	.65	.73	.30
4	.83	.68	.63	.72	.61	.45
5	.84	.70	.65	.63	.70	.44
6	.69	.68	.61	.64	.41	.53
7	.81	.72	.57	.56	.35	.48
8	.73	.70	.60	.59	.47	.59
9	.78	.76	.49	.52	.46	.64
10	.63	.75	.50	.57	.46	.45

S_1: Learning Culture
S_2: School/Family/Community Connections
S_3: Shared Leadership
S_4: Shared Goals for Learning
S_5: Purposeful Student Assessment
S_6: Effective Teaching

Table 11
AEL CSIQ Factor Loadings by Item Number and Factor Subscale, Combined Group

	Factor Number and Corresponding Subscale					
Item No.	1/S_3	2/S_6	3/S_4	4/S_1	5/S_2	6/S_5
1	.76	.61	.45	.54	.44	.41
2	.78	.72	.63	.55	.65	.47
3	.77	.70	.63	.71	.74	.43
4	.82	.68	.67	.60	.72	.54
5	.86	.70	.70	.71	.64	.53
6	.69	.72	.65	.45	.66	.60
7	.81	.76	.58	.36	.43	.56
8	.73	.73	.64	.52	.40	-.56
9	.77	.77	.54	.51	.37	.58
10	.61	.76	.55	.49	.40	.42

S_1: Learning Culture
S_2: School/Family/Community Connections
S_3: Shared Leadership
S_4: Shared Goals for Learning
S_5: Purposeful Student Assessment
S_6: Effective Teaching

The loadings within each factor were remarkably consistent, both in magnitude and direction. In the three tables, there were no loadings less than .30 and only two negative loadings, both in S_5. For all three analyses, the six factors extracted accounted for around 60 percent of the total variance.

A difference in the results of the factor analyses in Tables 9, 10, and 11 was the order in which subscales were associated with the factors extracted. No two orders matched exactly and differences and similarities can be seen by listing the orders.

Elementary	S_6	S_3	S_4	S_2	S_1	S_5
High School	S_3	S_6	S_4	S_2	S_1	S_5
Combined	S_3	S_6	S_4	S_1	S_2	S_5

Subscale 6, effective teaching, and Subscale 3, Shared Leadership, were associated with the first and second factors extracted for all three groups. Subscale 4, Shared Goals for Learning, was the third factor extracted for all three analyses, and Subscale 5, Purposeful Student Assessment, was associated with the final factor extracted, also for the three analyses. Finally, the first two subscales were in either the fourth or fifth position relative to extracted factors.

The three factor analyses provided statistical support for the construct validity of the AEL CSIQ. Given that six factors were quite clearly extracted and the subscales had clear associations with these factors supported the notion that six major constructs underlie continuous learning and improvement as measured by the AEL CSIQ. Too, the patterns of association of subscales with the order of the factors shows the similarity of results across the different groups.

Gender Effects

The respondents in the field test were identified by gender. However, gender is partially confounded with type of school because the gender split is not equal or proportional across type of school. For example, the majority of elementary teachers are female.

Table 12 shows the number of respondents and the subscale and total score means by gender. The frequencies show that the majority of the field-test respondents, about 80%, were female. Also, the females consistently had the higher mean scores on the AEL CSIQ subscales and total scores. To test for differences between the female and male respondents on the six subscales and the total score, independent t-tests were computed. The results of the t-tests showed significant difference, at the .001 level, in favor of the females on all of the subscales and total score except one (S_1, Shared Leadership).

Table 12

Frequencies (N) and Means (M) for AEL CSIQ Subscales and Total Score by Gender

Subscale/ Total	Female		Male	
	N	M	N	M
S_1	2909	48.4	684	46.1
S_2	2897	46.1	684	44.3
S_3	2836	46.7	677	46.0
S_4	2805	47.8	673	45.2
S_5	2827	47.3	668	44.5
S_6	2903	49.8	684	47.7
Total	2545	285.6	618	273.6

S_1: Learning Culture
S_2: School/Family/Community Connections
S_3: Shared Leadership
S_4: Shared Goals for Learning
S_5: Purposeful Student Assessment
S_6: Effective Teaching

CONCLUSIONS

The pilot and field tests have provided much information about the AEL CSIQ and its measurement. With data from over 3,800 educators in 132 schools, the statistics generated have high stability and precision. This section presents the conclusions drawn from the analyses, divided by testing periods.

Pilot Test

The following conclusions are warranted from the pilot study.

1. Although the original 147-item version of the AEL CSIQ had numerous items with defects in construction, those defects could be corrected readily, at least so that the 15-item and 12-item versions consisted entirely of adequately constructed items.

2. Internal consistency reliability of the AEL CSIQ is high and reliability is not affected in any noteworthy manner when the number of items is reduced.

3. The item-to-subscale correlations tend to be very stable across the various lengths of the AEL CSIQ subscales.

4. Relative positioning of the means is not affected by reducing the length of the AEL CSIQ.

5. The AEL CSIQ scale of measurement is appropriate in that the means are located so there is adequate room to measure change in either direction.

6. The stability of the measurement within AEL CSIQ subscales is very high across different numbers of items, as indicated by the very consistent coefficients of variation.

7. The relatively substantial correlations among the subscale scores indicate that the constructs of school improvement are interrelated.

8. Concurrent validity of the AEL CSIQ is high and it remains consistently high with 90 and 72 items in the total inventory.

Field Tests

The following conclusions are based on the results of the field tests.

1. Within a short period of time of about three weeks, the AEL CSIQ has moderate to high stability reliability. It should be noted that this result was based on a situation of no planned school improvement intervention.

2. The underlying constructs of continuous learning school improvement as measured by the six AEL CSIQ subscales account for almost two-thirds of the variance in the item scores, and for this type of measurement, two-thirds is a substantial portion.

3. Reducing the length of the AEL CSIQ to 10 items per subscale was very successful because it did not adversely affect reliability or validity and, at the same time, decreased the time needed for it to be completed by respondents.

4. Female respondents scored higher than male respondents on the subscales and total score, but this effect is partially confounded with type of school because the majority of educators in elementary schools are female.

In summary, the testing of the AEL CSIQ in the field has fostered its development into an excellent, usable 60-item inventory. The AEL CSIQ was reduced from 147 items with unequal numbers of items in subscales to a much more convenient and manageable length of 10 items in each subscale. More important, the AEL CSIQ was found to be very sound technically. Internal consistency reliability was high for all subscales, as was stability reliability. It showed satisfactory concurrent validity with another school climate instrument. The items had face validity based on item content, and the factor analyses provided strong support for the construct validity of the entire inventory as reflected by the subscales.

As used in these pilot and field tests, the AEL CSIQ was valid and reliable.

REFERENCES

Howley-Rowe, C. (2000a). *Lessons learned from the Quest for quality learning communities project, 1996-2000.* Charleston, WV: AEL.

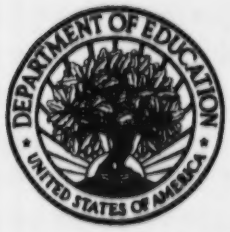
Howley-Rowe, C. (2000b). *Bending Knee Elementary: A case study of the Quest network.* Charleston, WV: AEL.

Howley-Rowe, C. (2000c). *Bowman Elementary: A case study of the Quest network.* Charleston, WV: AEL.

Howley-Rowe, C. (2000d). *Xavier High School: A case study of the Quest network.* Charleston, WV: AEL.

Manning, G., Curtis, K., & McMillen, S. (1996). *Building community: The human side of work.* Cincinnati, OH: Thompson Executive Press.

TM034423



U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

NOTICE

Reproduction Basis

This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").